

PCT09

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/623,514A

DATE: 08/03/2001
TIME: 14:00:18

Input Set : A:\43922new.txt
Output Set: N:\CRF3\08032001\I623514A.raw

ENTERED

3 <110> APPLICANT: National Research Council of Canada
4 Zou, Jitao
5 Taylor, David C
6 Wei, Yangdou
7 Jako, Colette C
9 <120> TITLE OF INVENTION: Diacylglycerol Acyl Transferase Gene from Plants
11 <130> FILE REFERENCE: 43922
W--> 12 <140> CURRENT APPLICATION NUMBER: US 09/623,514A
13 <141> CURRENT FILING DATE: 2000-10-03
14 <150> PRIOR APPLICATION NUMBER: PCT/CA99/01202
15 <151> PRIOR FILING DATE: 1999-12-16
16 <150> PRIOR APPLICATION NUMBER: US 60/112,812
17 <151> PRIOR FILING DATE: 1998-12-17
19 <160> NUMBER OF SEQ ID NOS: 25
21 <170> SOFTWARE: PatentIn Ver. 2.1

23 <210> SEQ ID NO: 1

24 <211> LENGTH: 1904

25 <212> TYPE: DNA

26 <213> ORGANISM: Arabidopsis thaliana

28 <400> SEQUENCE: 1

29 atttcttagc ttcttcttcc aatccgctct ttcctctctc attagattct gtttctcttt 60
30 tcaattttctt ctgcatgctt ctgattcttc tctgacgcct cttttctccc gacgctgttt 120
31 cgtcaaacgc ttttcgaaat ggcgattttg gattctgctg gcgttactac ggtgacggag 180
32 aacgggtggcg gagagttcgt cgatcttgat aggtctcgtc gacggaaatc gagatcggat 240
33 tcttctaacg gacttcttct ctctggttcc gataataatt ctcttcgga tgatgttga 300
34 gctcccgcgc acgttaggga tgggattgat tccgtgttta acgatgacgc tcagggaaca 360
35 gccaatattg cggagataa taacgggtgg ggcgataata acgggtgggtg aagaggcggc 420
36 ggagaaggaa gaggaacgc cgatgctacg tttacgtatc gaccgtcggg tccagctcat 480
37 cggagggcga gagagagtc acttagctcc gacgcaatct tcaaacagag ccatgccgga 540
38 ttattcaacc tctgtgtagt agttcttatt gctgtaaaaca gtagactcat catcgaaaat 600
39 cttatgaagt atggttggtt gatcagaacg gatttctggt ttagttcaag atcgtgcga 660
40 gattggccgc ttttcatgtg ttgtatatcc ctttcgatct ttcctttggc tgcctttacg 720
41 gttgagaaat tggtaactca gaaatacata tcagaacctg ttgtcatott tcttcatatt 780
42 attatcacca tgacagaggt tttgtatcca gtttacgtca ccctaagggtg tgattctgct 840
43 tttttatcag gtgtcacttt gatgctctc aottgcattg tgtggctaaa gttggtttct 900
44 tatgtcata ctagctatga cataagatcc cttagccaatg cagctgataa ggccaatcct 960
45 gaagtctctt actacgttag cttgaagagc ttggcatatt tcatggctgc tcccacattg 1020
46 tgttatcagc caagttatcc acgttctgca tgtatacggg aggggtgggt ggctcgtcaa 1080
47 tttgcaaaac tggatcatatt caccggattc atgggattta taatagaaca atatataaat 1140
48 cctattgtca ggaactcaaa gcatcctttg aaaggcgatc ttctatatgc tattgaaaga 1200
49 gtgttgaagc tttcagttcc aaatttatat gtgtggctct gcatgttcta ctgcttcttc 1260
50 cacctttggt taaacatatt ggcagagctt ctctgcttcg gggatcgtga attctacaaa 1320
51 gattggtgga atgcaaaaag tgtgggagat tactggagaa tgtggaatat gcctgttcat 1380
52 aaatggatgg ttcgacatat atacttcccg tgccttgcga gcaagatacc aaagacactc 1440
53 gccattatca ttgctttcct agtctctgca gtctttcatg agctatgcat cgcagttcct 1500
54 tgtcgtctct tcaagctatg ggcttttctt gggattatgt ttcaggtgcc tttggtcttc 1560
55 atcacaaact atctacagga aagggttggc tcaacgggtg ggaacatgat cttctggttc 1620

RAW SEQUENCE LISTING

DATE: 08/03/2001

PATENT APPLICATION: US/09/623,514A

TIME: 14:00:18

Input Set : A:\43922new.txt

Output Set: N:\CRF3\08032001\I623514A.raw

```

56 atcttctgca ttttcggaca accgatgtgt gtgcttcttt attaccacga cctgatgaac 1680
57 cgaaaaggat cgatgtcatg aaacaactgt tcaaaaaatg actttcttca aacatctatg 1740
58 gcctcgttgg atctccgttg atgttggtgt ggttctgatg ctaaaacgac aaatagtgtt 1800
59 ataaccattg aagaagaaaa gaaaattaga gttgttgat ctgcaaaaat tttggtagag 1860
60 acacgcaaac ccgtttggat tttgttatgg tgtaaagcgg ccgc 1904
63 <210> SEQ ID NO: 2
64 <211> LENGTH: 520
65 <212> TYPE: PRT
66 <213> ORGANISM: Arabidopsis thaliana
68 <400> SEQUENCE: 2
69 Met Ala Ile Leu Asp Ser Ala Gly Val Thr Thr Val Thr Glu Asn Gly
70 1 5 10 15
72 Gly Gly Glu Phe Val Asp Leu Asp Arg Leu Arg Arg Arg Lys Ser Arg
73 20 25 30
75 Ser Asp Ser Ser Asn Gly Leu Leu Ser Gly Ser Asp Asn Asn Ser
76 35 40 45
78 Pro Ser Asp Asp Val Gly Ala Pro Ala Asp Val Arg Asp Arg Ile Asp
79 50 55 60
81 Ser Val Val Asn Asp Asp Ala Gln Gly Thr Ala Asn Leu Ala Gly Asp
82 65 70 75 80
84 Asn Asn Gly Gly Gly Asp Asn Asn Gly Gly Gly Arg Gly Gly Gly Glu
85 85 90 95
87 Gly Arg Gly Asn Ala Asp Ala Thr Phe Thr Tyr Arg Pro Ser Val Pro
88 100 105 110
90 Ala His Arg Arg Ala Arg Glu Ser Pro Leu Ser Ser Asp Ala Ile Phe
91 115 120 125
93 Lys Gln Ser His Ala Gly Leu Phe Asn Leu Cys Val Val Val Leu Ile
94 130 135 140
96 Ala Val Asn Ser Arg Leu Ile Ile Glu Asn Leu Met Lys Tyr Gly Trp
97 145 150 155 160
99 Leu Ile Arg Thr Asp Phe Trp Phe Ser Ser Arg Ser Leu Arg Asp Trp
100 165 170 175
102 Pro Leu Phe Met Cys Cys Ile Ser Leu Ser Ile Phe Pro Leu Ala Ala
103 180 185 190
105 Phe Thr Val Glu Lys Leu Val Leu Gln Lys Tyr Ile Ser Glu Pro Val
106 195 200 205
108 Val Ile Phe Leu His Ile Ile Ile Thr Met Thr Glu Val Leu Tyr Pro
109 210 215 220
111 Val Tyr Val Thr Leu Arg Cys Asp Ser Ala Phe Leu Ser Gly Val Thr
112 225 230 235 240
114 Leu Met Leu Leu Thr Cys Ile Val Trp Leu Lys Leu Val Ser Tyr Ala
115 245 250 255
117 His Thr Ser Tyr Asp Ile Arg Ser Leu Ala Asn Ala Ala Asp Lys Ala
118 260 265 270
120 Asn Pro Glu Val Ser Tyr Tyr Val Ser Leu Lys Ser Leu Ala Tyr Phe
121 275 280 285
123 Met Val Ala Pro Thr Leu Cys Tyr Gln Pro Ser Tyr Pro Arg Ser Ala
124 290 295 300
126 Cys Ile Arg Lys Gly Trp Val Ala Arg Gln Phe Ala Lys Leu Val Ile

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/623,514A

DATE: 08/03/2001
TIME: 14:00:18

Input Set : A:\43922new.txt
Output Set: N:\CRF3\08032001\I623514A.raw

```

127 305          310          315          320
129 Phe Thr Gly Phe Met Gly Phe Ile Ile Glu Gln Tyr Ile Asn Pro Ile
130          325          330          335
132 Val Arg Asn Ser Lys His Pro Leu Lys Gly Asp Leu Leu Tyr Ala Ile
133          340          345          350
135 Glu Arg Val Leu Lys Leu Ser Val Pro Asn Leu Tyr Val Trp Leu Cys
136          355          360          365
138 Met Phe Tyr Cys Phe Phe His Leu Trp Leu Asn Ile Leu Ala Glu Leu
139          370          375          380
141 Leu Cys Phe Gly Asp Arg Glu Phe Tyr Lys Asp Trp Trp Asn Ala Lys
142 385          390          395          400
144 Ser Val Gly Asp Tyr Trp Arg Met Trp Asn Met Pro Val His Lys Trp
145          405          410          415
147 Met Val Arg His Ile Tyr Phe Pro Cys Leu Arg Ser Lys Ile Pro Lys
148          420          425          430
150 Thr Leu Ala Ile Ile Ile Ala Phe Leu Val Ser Ala Val Phe His Glu
151          435          440          445
153 Leu Cys Ile Ala Val Pro Cys Arg Leu Phe Lys Leu Trp Ala Phe Leu
154          450          455          460
156 Gly Ile Met Phe Gln Val Pro Leu Val Phe Ile Thr Asn Tyr Leu Gln
157 465          470          475          480
159 Glu Arg Phe Gly Ser Thr Val Gly Asn Met Ile Phe Trp Phe Ile Phe
160          485          490          495
162 Cys Ile Phe Gly Gln Pro Met Cys Val Leu Leu Tyr Tyr His Asp Leu
163          500          505          510
165 Met Asn Arg Lys Gly Ser Met Ser
166          515          520
170 <210> SEQ ID NO: 3
171 <211> LENGTH: 5193
172 <212> TYPE: DNA
173 <213> ORGANISM: Arabidopsis thaliana
175 <400> SEQUENCE: 3
176 gctcagcacc cattcttccc gttccatttg gttttattta ttccaaagtt taatattcct 60
177 tttgtataac attcaaactc tcacatgatt gattgtgtga aaaccccaca gattttacta 120
178 caataggggg agttgactta aaatagctat tgatgtcgaa aaaatgtatt ttagttataa 180
179 attatactaa agaaaatttt tgatttgtct gttgtttaag catatgtatt gttaaactta 240
180 aaaaaatatg tattgttaat cttaaaaatg taggagtaca catcaaatac tcgagcataa 300
181 tcaaaaccgt attcatagac cgatgtgaga atcaaataga agataatgtg atttttttaa 360
182 atatcgatc tccaaatcaa tcacttagaa gataatgtaa ttctttatgt gctacataaa 420
183 taaatatata tatatatata tatatatatc ttgtatatat gtcttgacaa aaaattgcca 480
184 gtcaaaaacc atgactgaat caaactataa gtcggattga atcaaactat aagtcggatg 540
185 agtattaatt tccattatgt ttctatactt tacaaaccgg aaaatagata ttatagatac 600
186 caaaaaagta gatttgtgta tattattaga agatttggaa ttatcatcatt atcaggatct 660
187 aaagtacttc cctaattaaa tcatgtcggt tgaaaaagct caatgaatgt ttgaaatttg 720
188 gaaagtttat taaattcgga tctttttttt ttgtttgtcg tcccaaacat ttttatttta 780
189 ttacaaataa tcaacttata cttactacta aatcatttca tatctttgat accaacaagt 840
190 catttcatat tctattttga tgtttaagaa aacactattt accagttaca aaatattata 900
191 aggattgttg tttagaaaaa aaagtacaag ttgaattctt tttgtcaa ataaaaattga 960
192 ctttttaata tataattgac ttattgaaca tgattacaga attaatcatc taaaaaactt 1020

```

RAW SEQUENCE LISTING

DATE: 08/03/2001

PATENT APPLICATION: US/09/623,514A

TIME: 14:00:18

Input Set : A:\43922new.txt

Output Set: N:\CRF3\08032001\I623514A.raw

```
193 tccaagttaa taataaatac atttcaaaga ctattagttc ttctttaaatt atttctaaaa 1080
194 gtgatcaaag actaccacat ataattcaga aaaagtagaa gttgatttct ttttgtcaaa 1140
195 taaataattg acttaaaata gtttggaag ccattgaact tgattataga attgataatg 1200
196 tacataaaaa aattccaagt ttataataaa tacatttttc aaatgctata tcagttcttc 1260
197 ttaaaatatt tcaactaaaa aacactcaaa tatagaataa atttattgaa taacatacca 1320
198 actgtaaaac agaatttgac aaaaaaaata aaaaaatgaa atgaagatga agacaaaaat 1380
199 aaatcaccag aggatcttat gcaaaaaaat atatgaatac acaataaacc atattgatat 1440
200 ttttaaaata aaataaaaac agaaaaatat cccaacaccg cttttcaatt aaaaatcttc 1500
201 cgtcaccatt gttgtcatct tctctctcgc tgaatccttt ttcttttctt cttcttcttc 1560
202 tcttcagaga aaactttgct tctctttcta taaggaacca gacacgaatc ccattcccac 1620
203 cgattttctta gcttcttctt tcaatccgct ctttccctct ccattagatt ctgtttcttc 1680
204 tttcaatttc ttctgcatgc ttctcgatcc tctctgacgc ctcttttctc ccgacgctgt 1740
205 ttogtcaaac gcttttcgaa atggcgattt tggattctgc tggcgttact acggtgacgg 1800
206 agaacggtgg cggagagttc gtcgatcttg ataggcttcg tcgacggaaa tcgagatcgg 1860
207 attcttctaa cggacttctt ctctctggtt ccgataataa ttctccttcg gatgatgttg 1920
208 gagctcccg cgcggttagg gatcggattg attccggttg taacgatgac gctcagggaa 1980
209 cagccaattt ggcgggagat aataacggtg gtggcgataa taacggtggt ggaagaggcg 2040
210 gcggagaagg aagaggaaac gccgatgcta cgtttacgta tcgaccgtcg gttccagctc 2100
211 atcggagggc gagagagagt ccacttagct ccgacgcaat cttcaaacag gtttaaaatc 2160
212 tcagaaatct tcgaatttgg tgtttgcttg ttgttttata tgggaattgag tttggtgatt 2220
213 gttttgcatt gcagagccat gccgattat tcaacctctg tgtagtagtt cttattgctg 2280
214 taaacagttag actcatcatc gaaaatctta tgaaggtttg ctgttacttg tttctccttt 2340
215 taggaattga attgcttgaa aatttatcag agacgaataa ctttggtgtt gctatcattc 2400
216 atgtagtatg gttggttgat cagaacggat ttctggttta gttcaagatc gctgcgagat 2460
217 tggccgcttt tcatgtgttg gtaaaagaag atgtttttta tttccagcaa tgttacattg 2520
218 ttatacgtat aatgatgagt ttagtgatca agttcctctt tgattcttct ttcttggtgc 2580
219 agtatatccc tttcgatctt tcttttggtt gcctttacgg ttgagaaatt ggtacttcag 2640
220 aaatacatat cagaacctgt gagtaattac tattctccag ccattactgt aatttttatt 2700
221 gaagacaagt ttgtatcatg aagaacttac aagttctggt ttgaaaatgc tcaaggttgt 2760
222 catctttctt catattatta tcaccatgac agaggttttg tatccagttt acgtcaccct 2820
223 aaggtgatac tgtttttctg gtctcagttt gtgatactgt ttttaagttt agttgtctga 2880
224 cccggtgatc ttgaaaatgg acaggtgtga ttctgctttt ttatcaggtg tcactttgat 2940
225 gctctcact tgcattgtgt ggctaaagt ggtttcttat gctcatacta gctatgacat 3000
226 aagatcccta gccaatgcag ctgataaggt aaaatacgaa aaagaagcgt atgtattagt 3060
227 cacttgcact gtgttactgt tttaaccaa cactgttatg aactttaggc caatcctgaa 3120
228 gtctcctact acgttagctt gaagagcttg gcataattca tggtogetcc cacattgtgt 3180
229 tatcaggtaa ctgcaaagtg catcaacctt tcttataact gcaagagttt cttgtctaaa 3240
230 cctcgatctt ttgcttttcc ccagccaagt tatccacgtt ctgcatgtat acggaagggt 3300
231 tgggtggctc gtcaatttgc aaaactggct atattcaccg gattcatggg atttataata 3360
232 gaacaagtac gttttcacat cttgctttat tagttttcct tggtgaaaat catcatccct 3420
233 gcgttgtcac cacttgactt catgttcttt tgttacattt tggcagtata taaatcctat 3480
234 tgtcaggaac tcaaagcatc ctttgaaagg cgatcttcta tatgctattg aaagagtgtt 3540
235 gaagctttca gttccaaatt tatatgtgtg gctctgcatg ttctactgct tcttccacct 3600
236 ttggtatgct gtgatcccat ctcttttcaa ataatttgca aattcgaaaa accgaaaaag 3660
237 gctaaatctc atacgaattt gatattttta gtttctttag gtcggtgatg taatttcagt 3720
238 tactgaacgc aaatctcttg tccaaagggt aaacatattg gcagagcttc tctgcttcgg 3780
239 ggatcgtgaa ttctacaaag attggtggaa tgcaaaaagt gtgggagatg tgagctatct 3840
240 tactcaaaag aaaacttatg atttttaatg ttgtcgttgt ttttgggtca tctaactaac 3900
241 caaatcatg tattcactgt cttcctttat cagtactgga gaatgtggaa tatggtatgg 3960
```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/623,514A

DATE: 08/03/2001

TIME: 14:00:18

Input Set : A:\43922new.txt

Output Set: N:\CRF3\08032001\I623514A.raw

```

242 ttctcttctt aaacatcacc ttcttttgta cacaaaatag aagaagagag ctaattaaga 4020
243 tcttggtttt cttgacagcc tggtcataaa tggatgggtc gacatatata cttcccggtg 4080
244 ttgcgagcga agataccaaa ggtgagtgag atatataccg atatgcaatt gtcgagattt 4140
245 gtttctgtga tataaattta accctccaca cacttgtttt tcagacactc gccattatca 4200
246 ttgctttcct agtctctgca gtctttcatg aggtatacat actttctaca ttgccctgtc 4260
247 tctagacgca tgaacacacg ctagtgaag aaatgctaatt attcaaagca ttgtttttac 4320
248 ttaacgatct tgtgttacaa atttcctttt gacagctatg catcgagtt ccttgctgtc 4380
249 tcttcaagct atgggctttt cttgggatta tgtttcaggt taaaaaatta ctaaactgct 4440
250 gcagtcgatt tttactaaac tctaattctca tattctgacc aaccaatttg tttgagtagg 4500
251 tgcctttggg cttcatcaca aactatctac aggaaagggt tggctcaacg gtatgctctc 4560
252 aaaacccgag aaaatagaac gaataactct ttctttcata gcctagccat ttaaactgca 4620
253 atgctgaaac ttaataataa aggtgatctg ttttggaatg ggatcatatt attaggtggg 4680
254 gaacatgac ttctgggtca tcttctgcat tttcggacaa ccgatgtgtg tgcttcttta 4740
255 ttaccacgac ctgatgaacc gaaaaggatc gatgtcatga aacaactggt caaaaaatga 4800
256 ctttcttcaa acatctatgg cctcgttgga tctccgttga tgttggtgtg gttctgatgc 4860
257 taaaacgaca aatagtgtta taaccattga agaagaaaag aaaattagag ttgttgatc 4920
258 tgcaaaaatt ttggtagaga cgcgcgaacc cgtttggtt ttgttatggt gtaaagaaat 4980
259 ttcaatcaaa aaactgttgt aataattgtt accaaaaaga aatgcttttc tggaaacgag 5040
260 gggaaaaata gtagttttgt taggttttac tgtttggacc aaatctagta aaaaactttt 5100
261 tgtaataagg aaaaaaaaag aacaaatgtg ataaatgcat ggggattgta tgaaaccttc 5160
262 caataaagtt gattggtgtg cccgttttgg gga 5193

```

265 <210> SEQ ID NO: 4

266 <211> LENGTH: 498

267 <212> TYPE: PRT

268 <213> ORGANISM: mouse

270 <400> SEQUENCE: 4

```

271 Met Gly Asp Arg Gly Gly Ala Gly Ser Ser Arg Arg Arg Arg Thr Gly
272 1 5 10 15
274 Ser Arg Val Ser Val Gln Gly Gly Ser Gly Pro Lys Val Glu Glu Asp
275 20 25 30
277 Glu Val Arg Asp Ala Ala Val Ser Pro Asp Leu Gly Ala Gly Gly Asp
278 35 40 45
280 Ala Pro Ala Pro Ala Pro Ala His Thr Arg Asp Lys Asp Gly
281 50 55 60
283 Arg Thr Ser Val Gly Asp Gly Tyr Trp Asp Leu Arg Cys His Arg Leu
284 65 70 75 80
286 Gln Asp Ser Leu Phe Ser Ser Asp Ser Gly Phe Ser Asn Tyr Arg Gly
287 85 90 95
289 Ile Leu Asn Trp Cys Val Val Met Leu Ile Leu Ser Asn Ala Arg Leu
290 100 105 110
292 Phe Leu Glu Asn Leu Ile Lys Tyr Gly Ile Leu Val Asp Pro Ile Gln
293 115 120 125
295 Val Val Ser Leu Phe Leu Lys Asp Pro Tyr Ser Trp Pro Ala Pro Cys
296 130 135 140
298 Val Ile Ile Ala Ser Asn Ile Phe Val Val Ala Ala Phe Gln Ile Glu
299 145 150 155 160
301 Lys Arg Leu Ala Val Gly Ala Leu Thr Glu Gln Met Gly Leu Leu Leu
302 165 170 175
304 His Val Val Asn Leu Ala Thr Ile Ile Cys Phe Pro Ala Ala Val Ala

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/623,514A

DATE: 08/03/2001

TIME: 14:00:19

Input Set : A:\43922new.txt

Output Set: N:\CRF3\08032001\I623514A.raw

L:12 M:283 W: Missing Blank Line separator, <140> field identifier
L:560 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:562 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:564 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:784 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10